

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number:

10/528,847

Source:

PCF

Date Processed by STIC:

01/17/2006

ENTERED

BEST AVAILABLE COPY



PCT

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/528,847

DATE: 01/17/2006

TIME: 13:46:45

Input Set : N:\DA\US10528847.raw

Output Set: N:\CRF4\01172006\J528847.raw

1 <110> APPLICANT: DSM IP ASSETS B.V.
2 <120> TITLE OF INVENTION: ACC gene
3 <130> FILE REFERENCE: NDR5217
4 <140> CURRENT APPLICATION NUMBER: US/10/528,847
5 <141> CURRENT FILING DATE: 2005-03-23
6 <150> PRIOR APPLICATION NUMBER: PCT/EP03/10683
7 <151> PRIOR FILING DATE: 2003-09-25
8 <150> PRIOR APPLICATION NUMBER: EP 02021625.5
9 <151> PRIOR FILING DATE: 2002-09-27
10 <160> NUMBER OF SEQ ID NOS: 10
11 <170> SOFTWARE: PatentIn version 3.2
13 <210> SEQ ID NO: 1
14 <211> LENGTH: 10561
15 <212> TYPE: DNA
16 <213> ORGANISM: Phaffia rhodozyma
17 <220> FEATURE:
18 <221> NAME/KEY: 5'UTR
19 <222> LOCATION: (1221)..(1222)
20 <220> FEATURE:
21 <221> NAME/KEY: exon
22 <222> LOCATION: (1446)..(1482)
23 <220> FEATURE:
24 <221> NAME/KEY: Intron
25 <222> LOCATION: (1483)..(1675)
26 <220> FEATURE:
27 <221> NAME/KEY: exon
28 <222> LOCATION: (1676)..(1758)
29 <220> FEATURE:
30 <221> NAME/KEY: Intron
31 <222> LOCATION: (1759)..(1832)
32 <220> FEATURE:
33 <221> NAME/KEY: exon
34 <222> LOCATION: (1833)..(1957)
35 <220> FEATURE:
36 <221> NAME/KEY: Intron
37 <222> LOCATION: (1958)..(2030)
38 <220> FEATURE:
39 <221> NAME/KEY: exon
40 <222> LOCATION: (2031)..(2171)
41 <220> FEATURE:
42 <221> NAME/KEY: Intron
43 <222> LOCATION: (2172)..(2243)
44 <220> FEATURE:

BEST AVAILABLE COPY

RAW SEQUENCE LISTING

DATE: 01/17/2006

PATENT APPLICATION: US/10/528,847

TIME: 13:46:45

Input Set : N:\DA\US10528847.raw

Output Set: N:\CRF4\01172006\J528847.raw

```

45 <221> NAME/KEY: exon
46 <222> LOCATION: (2244)..(2641)
47 <220> FEATURE:
48 <221> NAME/KEY: Intron
49 <222> LOCATION: (2642)..(2745)
50 <220> FEATURE:
51 <221> NAME/KEY: exon
52 <222> LOCATION: (2746)..(2991)
53 <220> FEATURE:
54 <221> NAME/KEY: Intron
55 <222> LOCATION: (2992)..(3074)
56 <220> FEATURE:
57 <221> NAME/KEY: exon
58 <222> LOCATION: (3075)..(3443)
59 <220> FEATURE:
60 <221> NAME/KEY: Intron
61 <222> LOCATION: (3444)..(3517)
62 <220> FEATURE:
63 <221> NAME/KEY: exon
64 <222> LOCATION: (3518)..(3552)
65 <220> FEATURE:
66 <221> NAME/KEY: Intron
67 <222> LOCATION: (3553)..(3625)
68 <220> FEATURE:
69 <221> NAME/KEY: exon
70 <222> LOCATION: (3626)..(3750)
71 <220> FEATURE:
72 <221> NAME/KEY: Intron
73 <222> LOCATION: (3751)..(3827)
74 <220> FEATURE:
75 <221> NAME/KEY: exon
76 <222> LOCATION: (3828)..(4026)
77 <220> FEATURE:
78 <221> NAME/KEY: Intron
79 <222> LOCATION: (4027)..(4095)
80 <220> FEATURE:
81 <221> NAME/KEY: exon
82 <222> LOCATION: (4096)..(4911)
83 <220> FEATURE:
84 <221> NAME/KEY: Intron
85 <222> LOCATION: (4912)..(4983)
86 <220> FEATURE:
87 <221> NAME/KEY: exon
88 <222> LOCATION: (4984)..(5384)
89 <220> FEATURE:
90 <221> NAME/KEY: Intron
91 <222> LOCATION: (5385)..(5455)
92 <220> FEATURE:
93 <221> NAME/KEY: exon

```

BEST AVAILABLE COPY

RAW SEQUENCE LISTING

DATE: 01/17/2006

PATENT APPLICATION: US/10/528,847

TIME: 13:46:45

Input Set : N:\DA\US10528847.raw

Output Set: N:\CRF4\01172006\J528847.raw

94 <222> LOCATION: (5456)..(5608)
 95 <220> FEATURE:
 96 <221> NAME/KEY: Intron
 97 <222> LOCATION: (5609)..(5673)
 98 <220> FEATURE:
 99 <221> NAME/KEY: exon

100 <222> LOCATION: (5674)..(5805)

101 <220> FEATURE:
 102 <221> NAME/KEY: Intron
 103 <222> LOCATION: (5806)..(5870)
 104 <220> FEATURE:

105 <221> NAME/KEY: exon
 106 <222> LOCATION: (5871)..(6832)
 107 <220> FEATURE:
 108 <221> NAME/KEY: Intron
 109 <222> LOCATION: (6833)..(6898)
 110 <220> FEATURE:
 111 <221> NAME/KEY: exon
 112 <222> LOCATION: (6899)..(6976)
 113 <220> FEATURE:
 114 <221> NAME/KEY: Intron
 115 <222> LOCATION: (6977)..(7047)
 116 <220> FEATURE:
 117 <221> NAME/KEY: exon
 118 <222> LOCATION: (7048)..(7227)
 119 <220> FEATURE:
 120 <221> NAME/KEY: Intron
 121 <222> LOCATION: (7228)..(7295)
 122 <220> FEATURE:
 123 <221> NAME/KEY: exon
 124 <222> LOCATION: (7296)..(9160)
 125 <220> FEATURE:
 126 <221> NAME/KEY: Intron
 127 <222> LOCATION: (9161)..(9230)
 128 <220> FEATURE:
 129 <221> NAME/KEY: exon
 130 <222> LOCATION: (9231)..(9530)
 131 <220> FEATURE:
 132 <221> NAME/KEY: polyA_site
 133 <222> LOCATION: (9813)..(9814)

134 <400> SEQUENCE: 1
 135 caacagacag acaagggaac ttacgtgtac atactgggtct ttccaatgtc gggcggtcga 60
 136 gattaactrag aacaatactt gacaatcgaa tctcttattc tgccctagtt gaaggcgtct 120
 137 gttcaaattg atcaagatct tccaatcatt gacatccagg tattcgcatc cgactctgct 180
 138 cgtatgtact gttecgattt tcttatggcc accagatttc aactctgata tacattgggt 240
 139 caccctgtct ttgtctcttt gcttttcgtt ccatctagcg ctgttcaacg gatcactcag 300
 140 tggccttgac tcaactccct ctggaaacgtg tgcccttatct caggttctga tttctcctca 360
 141 gccagtatgc gcacaaagca ggcgtcgtga ctttttgctc cataagacct ctcagcgggg 420
 142 aatatatgac actcatacat cgatagctcg tatgttttct ttgatcactt cctaaaatgt 480

BEST AVAILABLE COPY

RAW SEQUENCE LISTING

DATE: 01/17/2006

PATENT APPLICATION: US/10/528,847

TIME: 13:46:45

Input Set : N:\DA\US10528847.raw

Output Set: N:\CRF4\01172006\J528847.raw

143	aacggcaact gacattcaac atgatgogct ttcataagatc aactacttcc gactacgatg	540
144	accgttcttc tatacagccc agtcagctcg tcgacctcac ataaagtgac tgagaccgcg	600
145	atctcgaaca tcttattcct tccaccgtta gctgagaagt ggattacacc atcaatagaa	660
146	tcattcaacc cgttcttgcc tggactaatg cgtcaggagc tcttggataa aggagaaata	720
147	gctgagcaga ccattcacctt ggatgatgtc cgtctgtggc tgaactccgg aggtcgagtg	780
148	gcgtgctgca acgcacttcg aggaatttgg gaagtgaacc tcgtttggag tgataaatga	840
149	gattacgaaa gtctgttcga aacatccatg cttcatgata accgataacg cttaaatctt	900
150	gagagtgcgc acatcgatcg ccttttatat atgggttgg gaaacataa agtgttcata	960
151	gactattgtt catatctctt aaagtacaaa gacgcattcta accctaagcc tgaatgattg	1020
152	gcaaaatcct agtaagaccg tgaattccg aagaatacgc agttcattaa taaagatata	1080
153	gcttaggttaa gcagcgggtt ccccccaac caacctcatc cgaaattccc caggggggtt	1140
154	agattctcaa gcttttgat ccccatcccg taaagtggc cttaaacctt tcatctctac	1200
155	ttgttacttc ttttcttctt gacctcttc cccactccc tctattctc tgaacgaact	1260
156	cgctccctg tccatctact cttcttcggc tttcttttgg gtttttactt ttctcgttcc	1320
157	tctccatct ttcctctctt tttcgtatct gtgggtaact ttgcattcaa gggccctcac	1380
158	acataacctt atatccatct tctccattc acacacatct gtactcaacc aacaaagctc	1440
159	acaag atg gtt gtc gat cac gag agc gta agg cat ttc atc g	1482
160	Met Val Val Asp His Glu Ser Val Arg His Phe Ile	
161	1 5 10	
162	gtaagcgttc ttgttctttt ccttgtcttg ctcctgcat tttcttaaac gatctaggaa	1542
163	gagagggaaa ttacatctgg tcaattttcc gcgtctttt ccttggggac aaaagaatgc	1602
164	ctttctgtga tcggagatcg gttgctgac tcttttctt tgttctttt gctctttccc	1662
165	tcccttttac cag gt gga aac gca ctt gag aac gcc cct ccg tca agc	1710
W--> 166	Gly Gly Asn Ala Leu Glu Asn Ala Pro Pro Ser Ser	
167	15 20	
168	gtc acc gat ttc gtt aga agt caa gat ggt cac acg gtc atc acc aaa	1758
169	Val Thr Asp Phe Val Arg Ser Gln Asp Gly His Thr Val Ile Thr Lys	
170	25 30 35 40	
171	gtcagtaatt ttcatttttt ccttcacgta gcctcagggc caaggagcta aattgcttct	1818
172	gtatcatttc tcag gtc ctc att gcc aac aac gga atc gct gct gta aaa	1868
173	Val Leu Ile Ala Asn Asn Gly Ile Ala Ala Val Lys	
174	45 50	
175	gag atc cga tca gtt cgt aaa tgg gct tac gag acg ttt gga gat gag	1916
176	Glu Ile Arg Ser Val Arg Lys Trp Ala Tyr Glu Thr Phe Gly Asp Glu	
177	55 60 65	
178	cga gcc atc gaa ttt acg gta atg gcc act cca gaa gat tt	1957
179	Arg Ala Ile Glu Phe Thr Val Met Ala Thr Pro Glu Asp Leu	
180	70 75 80	
181	gttcgtacca atcacataag ctttccttga gtcagggaca tctcttaatt aattcaactt	2017
182	gagcgccata cag g aag gtg aac tgc gac tat att cga atg gct gat cga	2067
W--> 183	Lys Val Asn Cys Asp Tyr Ile Arg Met Ala Asp Arg	
184	85 90	
185	gtc gtc gaa gtt cct gga gga act aac aac aac aat cac tct aac gtc	2115
186	Val Val Glu Val Pro Gly Gly Thr Asn Asn Asn Asn His Ser Asn Val	
187	95 100 105 110	
188	gac ctc atc gtt gac att gcc gag cga ttc aat ata cat gct gtt tgg	2163
189	Asp Leu Ile Val Asp Ile Ala Glu Arg Phe Asn Ile His Ala Val Trp	
190	115 120 125	
191	gct gga tg gtaagtaaaa taggacctta acatgttggg agaagagtgt	2211

BEST AVAILABLE COPY

RAW SEQUENCE LISTING

DATE: 01/17/2006

PATENT APPLICATION: US/10/528,847

TIME: 13:46:45

Input Set : N:\DA\US10528847.raw

Output Set: N:\CRF4\01172006\J528847.raw

192 Ala Gly Trp
 193 ccacttaaac gcgcttttctt tccatccgac ag g ggt cac gct tcg gaa aac ccc 2265
 W--> 194 Gly His Ala Ser Glu Asn Pro
 195 130 135
 196 aga ctt ccc gag tct ctc gcc gcc tca aag aac aag atc gtc ttc att 2313
 197 Arg Leu Pro Glu Ser Leu Ala Ala Ser Lys Asn Lys Ile Val Phe Ile
 198 140 145 150

199 ggt cct ccc gga tcc gct atg cga tcc ctt gga gac aag att tet teg 2361
 200 Gly Pro Pro Gly Ser Ala Met Arg Ser Leu Gly Asp Lys Ile Ser Ser
 201 155 160 165
 202 acc atc gtt gcc cag tct gcc cag gtg ccg tgt atg gcc tgg tct gga 2409
 203 Thr Ile Val Ala Gln Ser Ala Gln Val Pro Cys Met Ala Trp Ser Gly

204 170 175 180
 205 tca ggc atc act gat aca gag ctc agc cct cag ggc ttc gtg act gtg 2457
 206 Ser Gly Ile Thr Asp Thr Glu Leu Ser Pro Gln Gly Phe Val Thr Val
 207 185 190 195 200
 208 ccc gat ggg cca tat cag gct gct tgt gta aag acg gtg gag gat ggt 2505
 209 Pro Asp Gly Pro Tyr Gln Ala Ala Cys Val Lys Thr Val Glu Asp Gly
 210 205 210 215
 211 ttg gtg cga gcc gag aag atc ggt ttg cca gtt atg atc aag gcc tct 2553
 212 Leu Val Arg Ala Glu Lys Ile Gly Leu Pro Val Met Ile Lys Ala Ser
 213 220 225 230
 214 gag gga gga gga gga aag ggt atc cga atg gtt cac agc atg gac aca 2601
 215 Glu Gly Gly Gly Lys Gly Ile Arg Met Val His Ser Met Asp Thr
 216 235 240 245
 217 ttc aag aac tcc tac aac tcc gtc gct tcc gag gtg cca g gtaagttcac 2651
 218 Phe Lys Asn Ser Tyr Asn Ser Val Ala Ser Glu Val Pro
 219 250 255 260
 220 tctgtttgac tggagatttg agcacaatct ctacatggg agttcaagaa ggaataccca 2711
 221 ctcatgaatt gacgactgcg ttcttgacct ctag ga tct ccg att ttc atc atg 2765
 222 Gly Ser Pro Ile Phe Ile Met
 223 265

224 gcc ttg gct gga tct gct cga cat ttg gag gtc cag ctc ctt gct gat 2813
 225 Ala Leu Ala Gly Ser Ala Arg His Leu Glu Val Gln Leu Leu Ala Asp
 226 270 275 280
 227 cag tac gga aac gct atc tct ttg ttc ggt cga gat tgc tct gtt cag 2861
 228 Gln Tyr Gly Asn Ala Ile Ser Leu Phe Gly Arg Asp Cys Ser Val Gln
 229 285 290 295 300
 230 cga cga cat cag aag atc att gag gag gct ccc gtc acg atc gct cgt 2909
 231 Arg Arg His Gln Lys Ile Ile Glu Glu Ala Pro Val Thr Ile Ala Arg
 232 305 310 315
 233 cca gag aga ttc gaa gag atg gag aag gct gct gtc agg ttg gcc aag 2957
 234 Pro Glu Arg Phe Glu Glu Met Glu Lys Ala Ala Val Arg Leu Ala Lys
 235 320 325 330
 236 tta gta gga tat gtc agt gcc ggt acc gtc gaa t gtaaggaaca 3001
 237 Leu Val Gly Tyr Val Ser Ala Gly Thr Val Glu
 238 335 340
 239 aacagctacc tctcattctg ttttttcgag atagtcgaact tacatcactt ttcttttgcc 3061
 240 ggattttctt tag ac ctc tac tct cac gcc gac gac tca ttc ttc ttc 3109

BEST AVAILABLE COPY

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/528,847

DATE: 01/17/2006
TIME: 13:46:46

Input Set : N:\DA\US10528847.raw
Output Set: N:\CRF4\01172006\J528847.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:4; N Pos. 6,9,15,18,21,24

Seq#:5; N Pos. 3,6,12,15,18,21,24

BEST AVAILABLE COPY

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/528,847

DATE: 01/17/2006

TIME: 13:46:46

Input Set : N:\DA\US10528847.raw

Output Set: N:\CRF4\01172006\J528847.raw

L:166 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 1
L:183 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 1
L:194 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 1
L:241 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 1
L:379 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 1
L:478 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 1
~~L:1426 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0~~
~~L:1463 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0~~

BEST AVAILABLE COPY